Rachel King

ECE 3522: Stochastic Processes in Signals and Systems

Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 1912

# Problem Statement

The purpose of this lab is to revisit concepts surrounding mean and variance calculations. Specifically, we will explore the error between the calculating the mean and variance over various spans of the signal. This is important as calculating these parameters may be more useful over a small chuck of a signal than the entire thing. Additionally, we will explore the error of estimating the PDF (through a histogram technique from previous assignments) and the actual PDF by calculating the mean squared error, something that we have also done in previous assignments.

# Approach and Results

*Part 1 – Mean and Variance Error*

First, we must create a random signal for which we will find the mean and variance of the entire thing:

%generate uniformly distributed data set, 10^6 points

D = rand(1, 10e3);

%calculate mean and variance of entire signal

mn = mean(D);

vr = var(D);

z = length(D);

I only used a signal 10^3 points long because 10^6 took too long from my MATALB to run, and this accomplished the same results. Next, I calculated the difference between the mean of the whole signal and the mean and variance over increasingly large segments of the random signal:

%parse signal and evaluate errors at increasing values

for i = 1:z

 %parse

 P = D(1:i);

 %mean and varience of parsed signal

 u = mean(P);

 s = var(P);

 %store errors in matricies

 Em(i) = (u-mn);

 Ev(i) = (s-vr);

end

The results of the error are displayed in figures 1 and 2.





# Part 2



Figure - True distribution



Figure - estimated distribution



Figure - Mean squarred error between the two

# MATLAB Code

%generate uniformly distributed data set, 10^6 points

D = rand(1, 10e3);

%calculate mean and varience of entire signal

mn = mean(D);

vr = var(D);

z = length(D);

%parse signal and evaluate errors at increasing values

for i = 1:z

 %parse

 P = D(1:i);

 %mean and varience of parsed signal

 u = mean(P);

 s = var(P);

 %store errors in matricies

 Em(i) = (u-mn);

 Ev(i) = (s-vr);

end

figure(1)

%plot

semilogx(Em);

xlabel('Number of Points (n)');

ylabel('Error of mean estimate');

title('Mean Error', 'fontweight', 'bold');

figure(2)

semilogx(Ev);

xlabel('Number of Points (n)');

ylabel('Error of varience estimate');

title('Variance Error', 'fontweight', 'bold');

Part 2

function CA\_6\_2

%generate array of random

F = rand(1, 10e6);

z = length(F);

%actual uniform distribution

xbins = (0:1/z:1);

y = length(xbins);

h = hist(F, xbins);

Preal = h;

%estimated uniform distribution

xbins2 = (0:.01:1);

y = length(xbins2);

h = hist(F, xbins2);

Pest = h/100000;

sum = 0;

sum2 = 0;

for j = 1:z

 sum = sum + Preal(j);

end

for j = 1:y

 sum2 = sum2 + Pest(j);

end

sum

sum2

q = length(Preal);

r = length(Pest);

%calculate mean squared error

figure(1)

plot(xbins2, Pest)

xlabel('Data Bins');

ylabel('Index');

title('Estimated Uniform Distribution')

figure(2)

plot(xbins, Preal)

xlabel('Data Bins');

ylabel('Index');

title('Actual Uniform Distribution')

for i = 1:101

 seg\_real = Pest(1:i);

 seg\_est = Preal(1:i);

 M(i) = mse(seg\_real, seg\_est);

end

M

figure(3)

semilogx(M);

xlabel('Number of Points (n)');

ylabel('Mean Squared Error');

title('Error between actual and estimated')

end

# Conclusions

In this assignment, we found the relationship between the number of points we have, and the distribution (pdf) of a data set which is randomly generated. For both parts, it was evident that, when more points were taken, the data converged on some value. For part one, we were asked to find the mean and variance error when we took larger and larger parts of a signal. For each case, the error between the two converged to 0 after a sufficient number of points. For my experiment, I decided to, instead of using the million points that were assigned, only use 10^4 points; this was because of the long run time of my computer, and I believe that the same results have been achieved using much less points and time. The second part of the problem had us estimate the pdf, and then find an actual distribution. In this case, the mean squared error for this did converge upon a value, but it was not zero, which leads me to believe that the findings are not correct. However, this lab was definitely an excellent exercise in getting to know random numbers more, as well as the impact that the number of numbers chosen has upon the problem.